Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
http://siba-ese.unisalento.it/index.php/ejasa/index
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v10n2p408

A wavelet based hybrid SARIMA-ETS model to forecast electricity consumption

By Deb, Chakrabarty

Published: 14 October 2017

Electronic Journal of Applied Statistical Analysis
Vol. 10, Issue 02, October 2017, 408-430
DOI: 10.1285/i20705948v10n2p408

A wavelet based hybrid SARIMA-ETS model to forecast electricity consumption

Mallika Deb*and Tapan Kumar Chakrabarty

Department of Statistics, North-Eastern Hill University, East Khasi Hills, Shillong-793022, Meghalaya, India

Published: 14 October 2017

Energy plays a fundamental role in the process of economic growth of a nation. Growth in the demand and consumption of energy is linked to economic output of a country as measured by gross domestic product. In view of formulating sustainable strategies in the energy service market, modelling and forecasting future demand of electricity is an integral part of decision support system of energy production in developed and developing world. Recent years have witnessed an increasing interest in providing prediction models of electrical energy consumption with greater accuracy. Besides the use of logistic and Harvey logistic growth curve models, the stochastic forecasting has been carried out using Box-Jenkins ARIMA, Holt-Winter, SARIMA and time series ANN models etc. However, their performance is far from perfect and it is especially true when the data contain complex nonlinear pattern and volatility. In this article, we propose a hybrid model that splits a time series into an approximate and a detailed component via discrete wavelet transform and then SARIMA and ETS models are used to fit and forecast the wavelet approximation and the detailed component respectively. Using the real monthly electricity consumption data from eight north eastern states of India during 2004-2015, we have developed the proposed model. Results of our investigation successfully demonstrates the higher degree of prediction accuracy of the proposed model than a data driven Box-Jenkins ARIMA model in terms of various performance accuracy measurement statistics and produces a substantial reduction in forecast errors.

*Corresponding author: mallikadeb706@gmail.com

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

keywords: Discrete wavelet transform, Box-Jenkins ARIMA, SARIMA, ETS, Electricity consumption forecasting.

1 Introduction

Energy is one of the most essential ingredients for the economic growth and all round developments of any nation. Electricity being a highly versatile form of energy boosts the efficiency of performance in every sector of an economy. Growth in the demand and consumption of energy is linked to economic output of a country as measured by gross domestic product (GDP). During the recent years, the world has witnessed a significant increase in electricity demand due to increase in population and many other factors. It is vital for a country to be able to supply the electricity as per the demand in order to keep up the pace of development. If the production of electricity is less than the demand, the electricity dependent industries are affected and the growth declines; on the other hand, higher productions of electricity leads to the power plants remain idle and which is a waste in economic resources. Electricity demand forecasting is thus a key task in the planning and operation of power systems and electricity markets. There are numerous benefits of electricity forecasting in terms of optimum utilization of energy resources, improve stability of the system, improve availability, improve operation both from technical and economic considerations, improve quality of supply, improve grid discipline, improve service to an electricity-deficit area from an electricity-surplus area, coordinated planning for both maintenance and future growth of the system. Thus, an accurate and precise electricity demand forecasting method is an integral part of decision support system of energy production in developed and developing world.

Literature on forecasting electricity consumption has grown with numerous methodological and modelling choices, since, determining the future is very complex and tied with infinite possibilities. In forecasting an economic variable, past observations are used to predict the future values (Brockwell and Davis, 2006; Granger and Newbold, 2014; Wheelwright et al., 1998). Although there are arguments in against and in favour of using forecasts for policy analysis (Sims et al., 1986; Zahan and Kenett, 2013), many researchers support it arguing that forecasting provides guidelines for policymakers to take steps for the future based on past experiences. Although it is impossible to predict a future scenario exactly as it would be, but too many things run smooth over time. Therefore, researchers always strive to tackle this issue with minimum possible errors.

Over the years, a number of growth curve based models have been proposed to forecast electricity in the literature. Mohamed and Bodger (2005) investigated the effectiveness of two forms of Harvey models and a Logistic model for forecasting electricity consumption in New Zealand. The Logistic model is a time series extrapolation model. The model involves obtaining the saturation level of the electricity consumption using a Fibonacci search technique (Mohamed and Bodger, 2003). The estimated saturation level is used as a constant variable to fit the best logistic growth curve and forecasts are obtained by extrapolating the fitted curve. A major finding was that the Logistic model very effectively described the historical electricity consumption while producing lower forecasts in general. This is mainly due to the constraints imposed by the saturation level of the logistic growth curve. This opened the opportunity to search for other models that may give rise to higher forecasts.

A number of forecasting models for electricity consumption have been developed in the multiple linear regression framework using economic, social, geographic and demographic factors (Egelioglu et al., 2001; Yan, 1998; Mohamed and Bodger, 2003). Such models target to predict future growth or decline in the consumption pattern based on an assumed linear relationship among the response and the potential covariates influencing the response. Regression models may not be useful or result in very poor predictions when the response variable is not second order stationary. To overcome such situations, time series based methods like Box-Jenkins Auto-Regressive Integrated Moving Average (ARIMA), Holt-winter model, Seasonal ARIMA (SARIMA), Auto-Regressive Conditional Heteroscedasticity (ARCH), Generalised ARCH (GARCH), Dynamic Choice Artificial Neural Network (DCANN) etc. have also been used successfully by many researchers.

ARIMA model projects a time series univariate variable based on three parameters i.e. order of the moving average, order of differencing and order of autoregressive scheme. Erdogdu (2007), Ediger and Akar (2007) used ARIMA models to forecast electricity demand and primary energy demand by fuel in Turkey respectively. Taylor (2003) predicted the electricity demand for UK using the Holt-Winter model and concluded that the Holt-Winter model outperform as compared to well-fitted ARIMA models. The Holt-Winter forecasting model not only entertains the exponentially smoothed component but also the trend component while projecting a time series variable. Wang et al. (2012) proposed SARIMA to forecast electricity demand in China, while Garcia et al. (2005) used a GARCH forecasting model to predict day-ahead electricity prices. A novel model, known as, DCANN used by Wang et al. (2016) for an electricity price forecasting system.

The concept of combining methods was first introduced by Bates and Granger (1969), who proved that the combined methods were more efficient and easier than the individual ones as the hybrid models combine strengths of few individual models to render better prediction accuracy. Tan et al. (2010) developed a combined method by using three individual methods namely Wavelet Transform (WT), ARIMA and GARCH to forecast the electricity price. Liu and Shi (2013) applied various Auto-Regressive Moving Average (ARMA) models with GARCH processes, along with their modified forms, ARMA-GARCH-in-mean, to address the issue of forecasting hour-ahead electricity prices. Yan and Chowdhury (2013) presented a hybrid mid-term electricity market clearing price (MCP) forecasting model combining both least squares support vector machine (LSSVM) and ARMA with external input (ARMAX) modules. In (dos Santos Coelho and Santos, 2011), a Radial Basis Function-Neural Network-GARCH (RBF-NN-GARCH) model was proposed where the traditional RBF-NN model is extended by using GARCH specifications for modeling the variability of price signals.

It is argued that electricity consumption is affected by several periodical as well as time varying complex non-linear components other than simple trend that can be illustrated by using certain mathematical transformations. To improve the forecasting precision, the optimal forecasting model should make a wise use of these different data components.

Thus, we can decompose a given time series into component time series exhibiting typical time series features. Based on the features of these components so derived, we apply the appropriate forecasting scheme to each of these component time series and then we convert the forecasting of the time series at component levels back to the original time series level using the inverse transformation. One way of conducting this decomposition is via the WT. A wavelet is a mathematical function used to divide a given function or continuous-time signal into multi scale components.

In recent years, there is a tremendous surge in the use of WT based methods for their encouraging results in many domains of science, engineering, signal processing, and statistical problems including the short term electricity price forecasting (Zhang and Tan, 2013; Catalão et al., 2011; Shrivastava and Panigrahi, 2014; Afanasyev and Fedorova, 2016). However, the use of WT based models in electricity consumption has not yet been seen in literature so far. This article suggests a new potentially efficient hybrid technique based on discrete wavelet transform (DWT) invented by Haar (1910), which transforms a time series into two components; one is the approximation of the original series and the other is the detailed component. In order to improve the forecasting precision, the forecast models should be tailored for forecasting the components separately. After a prior decomposition of the series through 1-level DWT, this proposed approach adopts SARIMA and exponential smoothing state space (ETS) models to fit and forecast the wavelet approximation and the detailed component respectively. Thus, the proposed hybrid model tactically utilizes the unique strengths of those individual univariate methods mentioned above claiming to have a higher forecasting accuracy. The superiority in terms of forecast accuracy of the method is demonstrated with the real monthly electricity consumption data of North-Eastern Region (NER) of India with effect from April-2004 to December-2015 containing 141 observations and its forecasting accuracy is compared with that of SARIMA and ETS models. Our investigation successfully indicates that the proposed method is one of the most suitable electricity consumption forecasting techniques with the higher degree of prediction accuracy than the existing benchmark methods available in the literature.

The rest of the article is organised in the following manner. The proposed approach is discussed in detail in section 2. Section 3 describes the data used in this study and results of this investigation. The article ends with the section 4 as concluding remarks.

2 Proposed hybrid model

2.1 Discrete wavelet transform

The DWT is a linear transformation that operates on a data vector whose length is an integer power of two, transforming it into a numerically different vector of the same length. It is a tool that separates data into different frequency components, and then studies each component with resolution matched to its scale. The Haar transform (Haar, 1910) is performed in several stages or levels. The first level or 1-level is the mapping H_1 defined by

$$f \xrightarrow{H_1} (a^1 | d^1)$$

from a discrete signal f to its first trend a^1 and first fluctuation d^1 referred to as approximate and detailed components throughout this article. Generally we express a discrete signal in the form $f = (f_1, f_2, \ldots, f_N)$, where N is a positive even integer which is referred to as the length of f. The values of f are the N real numbers $f_1, f_2, \ldots, f_N$. These values are typically measured values of an analog signal g, measured at the time values $t = t_1, t_2, \ldots, t_N$. That is, the values of f are

$$f_1 = g(t_1), f_2 = g(t_2), \ldots, f_N = g(t_N)$$

For computation of the approximate and detailed components the reader is referred to the texts of Walker James (1999) and Gençay et al. (2001). In this article, however, we performed these computations using the package 'waveslim' (Whitcher, 2015) of statistical software 'R-3.3.0' (R Development Core Team, 2008). This function performs a level J decomposition of the input vector using the non-decimated discrete wavelet transform.

2.2 SARIMA model

In 1970, Box and Jenkins (Box and Jenkins, 1970), made ARIMA (Box et al., 1994) models popular by proposing a model building methodology comprising several stages: specification, estimation, diagnostic checking and forecasting. ARIMA is a widely used time series modelling technique. It uses historical time series patterns and therefore, does not require the dependent variable; instead, time series information is used to generate the series itself. Therefore, we explain the core of the ARIMA model here. The general $ARIMA(p, d, q)$ model is formulated as follows. Let the variable y_t denote the data value at any given time t,

$$\phi_p(B)\Delta^d y_t = \delta + \theta_q(B)\epsilon_t; t = 1, 2 \ldots n$$

where, $\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$ is the Auto-Regressive (AR) operator, $\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q$ is the Moving Average (MA) operator, $By_{t+1} = y_t$, Δ is the difference operator, δ is some drift, B is the time lag operator or backward shift operator, p is the order of AR operator, q is the order of MA operator and d is the order of the differencing operator and ϵ_t is a random error term which is Gaussian distributed with zero mean and constant variance σ^2 (i.e. a white noise process).

ARIMA model also may include nonstationary seasonal terms into the model. Seasonality in a time series is a regular seasonal pattern of changes that repeats over time span's', where's' defines the number of time periods until the pattern repeats again. A seasonal ARIMA model is formed by including additional seasonal terms in the ARIMA model denoted as $ARIMA(p, d, q)(P, D, Q)_s$; where s is the season length, P is the order of seasonal AR operator, Q is the order of seasonal MA operator and D is the order of the seasonal differencing operator. Seasonal $ARIMA(p, d, q)(P, D, Q)_s$ is defined as,

$$\phi_p(B)\Phi_P(B^s)\Delta^d\Delta_s^D y_t = \delta + \theta_q(B)\Theta_Q(B^s)\epsilon_t; t = 1, 2 \ldots n$$

where, $\Phi_P(B^s) = 1 - \Phi_s B^s - \Phi_{2s} B^{2s} - \cdots - \Phi_{Ps} B^{Ps}$ is the AR operator, $\Theta_Q(B^s) = 1 - \Theta_s B^s - \Theta_{2s} B^{2s} - \cdots - \Theta_{Qs} B^{Qs}$ is the MA operator.

The first step in modeling time index data is to convert the non-stationary time series to stationary one. This is important for the fact that a lot of statistical and econometric methods are based on this assumption and can only be applied to stationary time series. Non-stationary time series are erratic and unpredictable while stationary process is mean-reverting, i.e., it fluctuates around a constant mean with constant variance. SARIMA model is a linear model that represent both stationary and non-stationary data. In this technique, the given time series data are first checked for stationarity. Stationarity is typically expressed by requiring the AR and MA polynomials to have their roots outside the unit circle (Meitz and Saikkonen, 2013). After the seasonal and non-seasonal differencing, the data becomes stationary and the resultant data can be modeled as an ARMA time series as follows. The data value y_t at any given time t, is considered as a function of the previous p data values, say$y_{t-1}, y_{t-2} \dots y_{t-p}$ and the errors at times t, t-1 ... t-q, say $\epsilon_t, \epsilon_{t-1} \dots \epsilon_{t-q}$.

$$y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} \dots \phi_p y_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} \dots \theta_q \epsilon_{t-q}; t = 1, 2 \dots n$$

where, $\phi_1, \phi_2 \dots \phi_p$ are the AR coefficients and $\theta_1, \theta_2 \dots \theta_q$ are the MA coefficients.

Thus, the time series model described above is denoted as ARMA (p, q). Now, identifying the orders p, q is done using correlation analysis (Box and Jenkins, 1990), using the nature of the autocorrelation function and the partial autocorrelation function. The model coefficients are estimated using the Box-Jenkins method (Box and Jenkins, 1990).

2.3 ETS model

Forecasting methods based on exponential smoothing (ES) techniques proposed in the late 1950s (Brown, 1959; Holt, 1957, Winters, 1960) generate reliable forecasts for a wide spectrum of time series in industry and economics. Although these methods have been employed for many decades, recent methodological developments have embedded these models in modern dynamic nonlinear model framework. Hyndman et al. (2002, 2008) outline the ETS (Error-Trend-Seasonal or ExponenTial Smoothing) framework which defines an extended class of ES methods and offers a theoretical foundation for analysis of these models using state-space based likelihood calculations, with support for model selection and calculation of forecast errors.

Each exponential smoothing state space model consists of a measurement equation that describes the observed data and some transition equations that describe how the unobserved components or states (level, trend, seasonal) change over time. Hence these are referred to as 'state space models'. For each method there exist two models: one with additive errors and one with multiplicative errors. We label each state space model as ETS(*,*,*) for $(Error, Trend, Seasonal)$. The possibilities for each component are: Error = $\{A, M\}$, Trend = $\{N, A, A_d, M, M_d\}$ and Seasonal = $\{N, A, M\}$ where "N" stands for none, "A" stands for additive, "M" stands for multiplicative, "A_d" stands for additive damped and "M_d" stands for multiplicative damped. Therefore, in total there exist 30 such state space models: 15 with additive errors and 15 with multiplicative errors. The models considered in this paper are ETS(A,N,A) model which includes an additive seasonal component with additive errors and ETS(A,A,A) model which includes an additive trend and additive seasonal component both with additive errors.

Let the sequence $\{y_t; t = 1, 2 \dots n\}$ denote the observed data and $x_t = (l_t, b_t, s_t, s_{t-1} \dots s_{t-(m-1)}$ denote the unobserved state, where l_t denotes the level, b_t denotes the trend and s_t denotes the seasonal component, at time t. Then the models in subsection 2.3.1 and subsection 2.3.2 can be written in the form as equation (1) and equation (2).

$$y_t = h(x_{t-1}) + k(x_{t-1})\epsilon_t \tag{1}$$

$$x_t = f(x_{t-1}) + g(x_{t-1})\epsilon_t \tag{2}$$

where ϵ_t is a Gaussian white noise process with mean zero and variance σ^2.

2.3.1 ETS(A,N,A) model

ETS(A,N,A) model can be expressed as the sum of two components: viz, level and seasonal component. Mathematically, the model having the observation equation is defined as,

$$y_t = l_{t-1} + s_{t-m} + \epsilon_t \; ; \; \epsilon_t = y_t - l_{t-1} - s_{t-m} \sim NID(0, \sigma^2)$$

with m being the period of seasonality.

Level smoothing equation

$$l_t = \alpha(y_t - s_{t-m}) + (1-\alpha)l_{t-1}; 0 \leq \alpha \leq 1.$$

Seasonal smoothing equation

$$s_t = \gamma(y_t - l_{t-1}) + (1-\gamma)s_{t-m}; 0 \leq \gamma \leq 1.$$

The state space form of the ETS(A,N,A) model as well as its components are defined below.

Observation equation

$$y_t = l_{t-1} + s_{t-m} + e_t$$

since $e_t = \epsilon_t$, the residual, for additive error model.

Level equation

$$\begin{aligned} l_t &= \alpha(y_t - s_{t-m}) + (1-\alpha)l_{t-1} \\ &= l_{t-1} + \alpha(y_t - l_{t-1} - s_{t-m}) \\ &= l_{t-1} + \alpha e_t. \end{aligned}$$

Seasonal equation

$$\begin{aligned} s_t &= \gamma(y_t - l_{t-1}) + (1-\gamma)s_{t-m} \\ &= s_{t-m} + \gamma(y_t - l_{t-1} - s_{t-m}) \\ &= s_{t-m} + \gamma e_t. \end{aligned}$$

State equations $l_t = l_{t-1} + \alpha e_t$ and $s_t = s_{t-m} + \gamma e_t$.

2.3.2 ETS(A,A,A) model

ETS(A,A,A) model can be expressed as the sum of three components: viz, level, trend and seasonal component. Mathematically, the model having the observation equation is defined as,

$$y_t = l_{t-1} + b_{t-1} + s_{t-m} + \epsilon_t \ ; \ \epsilon_t = y_t - l_{t-1} - b_{t-1} - s_{t-m} \sim NID(0, \sigma^2).$$

Level smoothing equation

$$l_t = \alpha(y_t - s_{t-m}) + (1-\alpha)(l_{t-1} + b_{t-1}); 0 \leq \alpha \leq 1.$$

Trend equation

$$b_t = \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1}; 0 \leq \beta \leq 1.$$

Seasonal smoothing equation

$$s_t = \gamma(y_t - l_{t-1} - b_{t-1}) + (1-\gamma)s_{t-m}; 0 \leq \gamma \leq 1.$$

The state space form of the ETS(A,A,A) model as well as its components are defined below.

Observation equation

$$y_t = l_{t-1} + b_{t-1} + s_{t-m} + e_t.$$

Level equation

$$\begin{aligned} l_t &= \alpha(y_t - s_{t-m}) + (1-\alpha)(l_{t-1} + b_{t-1}) \\ &= l_{t-1} + b_{t-1} + \alpha(y_t - l_{t-1} - b_{t-1} - s_{t-m} \\ &= l_{t-1} + b_{t-1} + \alpha e_t. \end{aligned}$$

Trend equation

$$\begin{aligned} b_t &= \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1} \\ &= b_{t-1} + \beta(y_t - l_{t-1} - b_{t-1}) \\ &= b_{t-1} + \alpha\beta e_t. \end{aligned}$$

Seasonal equation

$$\begin{aligned} s_t &= \gamma(y_t - l_{t-1} - b_{t-1}) + (1-\gamma)s_{t-m} \\ &= s_{t-m} + \gamma(y_t - l_{t-1} - b_{t-1} - s_{t-m}) \\ &= s_{t-m} + \gamma e_t. \end{aligned}$$

State equations $l_t = l_{t-1} + b_{t-1} + \alpha e_t$, $b_t = b_{t-1} + \alpha\beta e_t$ and $s_t = s_{t-m} + \gamma e_t$.

2.4 Evaluation methods of model adequacy and performance accuracy measures

The evaluation methods of model adequacy provide critical guidance to the appropriate choice of models. The model validation is usually done by finding the Akaike's 'An Information Criterion' (AIC) (Akaike, 1974), Akaike's corrected 'An Information Criterion' (AICc) (Cavanaugh, 1997) or Schwarz's 'Bayesian Information criterion' (BIC) (Schwarz et al., 1978). The Akaike information criterion (AIC) is a measure of the relative quality of statistical models for a given set of data. Given a collection of models for the data, AIC estimates the quality of each model, relative to each of the other models. Hence, AIC provides a means for model selection. Suppose that we have a statistical model of some data. Let L be the maximum value of the likelihood function for the model; let k be the number of estimated parameters in the model. Then the AIC value of the model is AIC = 2k - 2ln(L) (Akaike, 1974; Burnham and Anderson, 2002) and AICc is AIC with a correction for finite sample sizes. The formula for AICc is, AICc = AIC $+\frac{2k(k+1)}{n-k-1}$ (Burnham and Anderson, 2002; Cavanaugh, 1997); where n denotes the sample size and k denotes the number of parameters. In this article, AICc values has been used for model selection. Given a set of candidate models for the data, the preferred model is the one with the minimum AICc value.

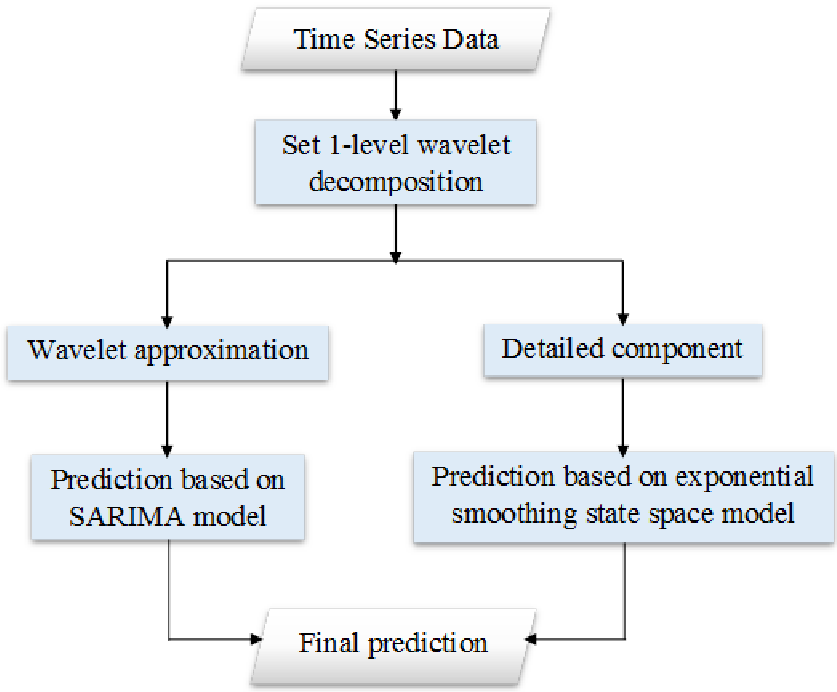

Figure 1: Proposed wavelet based hybrid SARIMA-ETS model

The steps of the algorithm for the proposed approach are represented as a flow chart in fig. 1. The proposed algorithm of wavelet based SARIMA-ETS model can give better accuracy as compared to using SARIMA or ETS model directly to the observed time series data without performing any modification of the original data. This fact has been verified using a real monthly electricity consumption data set.

All three methods: SARIMA, ETS and the proposed wavelet based hybrid SARIMA-ETS models are considered in this paper. A good forecasting method needs to take into consideration the degree of accuracy. So, performance accuracy measurement statistics (Hyndman and Koehler, 2006) used for comparison of prediction accuracy will be discussed in the present article. Several performance measures are considered as performance measure for accuracy comparison over the forecast horizon. The smaller the measure, the better the model.

Let, y_t denote the observation at time t, f_t denote the forecast of y_t and m is the number of observations in the forecast horizon. Then define the forecast error as $e_t = y_t - f_t$; t = 1, 2 . . . m and the percentage error as $p_t = 100e_t/y_t$; t = 1, 2 . . . m. The six measures used in the same prediction horizon. The measures are,

Scale-dependent measures:

Mean squared error (MSE) = $mean(e_t^2)$

Mean Absolute error (MAE) = $mean|e_t|$

Measures based on percentage errors:

Mean absolute percentage error (MAPE) = $mean|p_t|$

Root mean square percentage error (RMSPE) = $\sqrt{mean(p_t)^2}$

Symmetric mean absolute percentage error (sMAPE) = $mean\frac{2|y_t-f_t|}{|y_t+f_t|}$

An alternative way of scaling is to divide each measure by the measure obtained from the benchmark method used in the article. For instance, a relative MSE is given by, RelMSE = MSE/MSE_b; where MSE_b denote the MSE from the benchmark method. Similar measures can be defined using MAE, MAPE etc.

3 Data and Results

3.1 Data Description

Time series data on monthly electricity consumption of NER of India in megawatt (MW) unit during April-2004 through December-2015 has been used in this study. The data has been taken from the Annual Reports of North-Eastern Regional Power Committee (NERPC, 2015). Statistical reports published by NERPC have been providing monthly electricity data on a regular basis. These include data on electricity peak demand, electricity consumption, energy requirement, energy availability etc. for NER as a whole and also separately for each of the seven north-eastern states of India, namely, Arunachal Pradesh, Assam, Manipur, Meghalaya, Mizoram, Nagaland and Tripura.

3.2 Analysis and Results

The present analysis has been carried out using observed monthly electricity consumption containing 141 observations in MW provided in fig. 2. In order to do out-of-sample one-step forecast, two forecast horizons are considered in this article viz. forecast horizons of 31 and 48 observations. So the observed data is split into two parts, out of 141 observations first 110, 93 points are taken as training sets and the rest 31, 48 observations are taken as experimental sets or forecast horizons respectively.

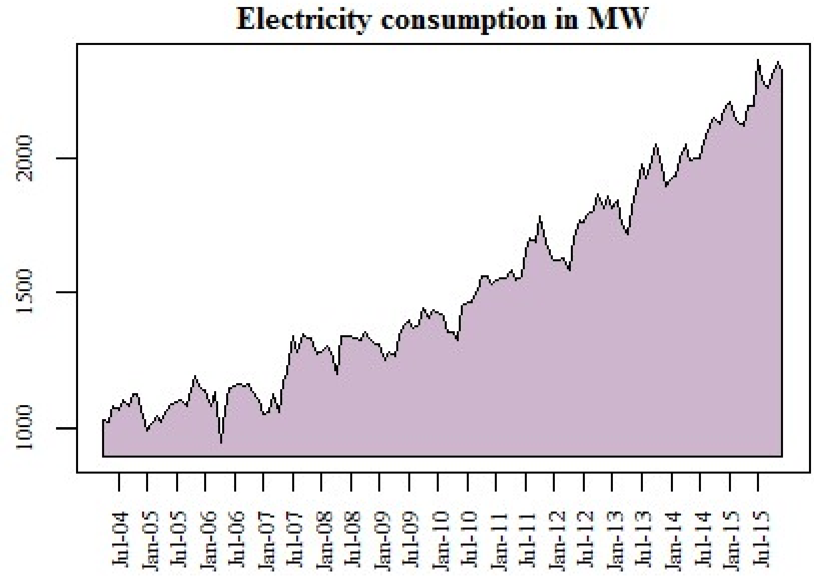

Figure 2: Observed monthly electricity consumption

Sample autocorrelation function (ACF) and sample partial autocorrelation function (PACF) were used to assess the time series behaviour viz. stationarity, seasonality of training sets of electricity consumption. Persistent patterns of slowly declining ACF for both training sets of electricity consumption were evident and thus indicating lack of stationarity but those do not indicate the existence of seasonality. One can formally test the significance of the seasonal component as follows. According to Rob Hyndman, there is a related log-likelihood test based on the difference between the selected model and the equivalent model with an additional seasonal term added. Twice the difference between the two log-likelihoods will have a chi-squared distribution according to Wilks' theorem. The degrees of freedom will be the difference in the number of parameters being estimated in the two models. If the hypothesis test is significant, we can conclude that the data are very unlikely to have been generated from the simpler (non-seasonal) model. The test results with p-value 1.873888e-09 and 1.795604e-06 show the additional seasonal component is significant. Also, AICc values support the significance of seasonality.

Then we have used the 'forecast' (Hyndman et al., 2014) package of statistical software 'R' for the purpose of fitting and forecasting different plausible models. Various univariate time series models were fitted viz. ARIMA model, linear model with time series components, ETS model, Holt-Winters time series model, structural Time Series model, cubic spline and theta method. Based on forecast accuracy, two higher accuracy performed models, ARIMA and ETS models are considered. So, an ETS model with additive trend and additive seasonal component ETS(A,A,A) model along with additive errors has been applied to the training set of electricity consumption with 110 observations and a model with additive seasonal component ETS(A,N,A) model including additive errors has been applied to the training set of electricity consumption with 93 observations. Based on the AICc values, the $ARIMA(1,1,0)(2,0,0)_{12}$ model was fitted to both training sets of electricity consumption, indicating a first order non-seasonal difference, and non-seasonal AR(1) and seasonal AR(2) components. The model descriptions are given in the table 1 and table 2 which provide the estimates of the coefficients along with standard errors and AICc values for fitted models.

In order to check out any scope of improvement in the prediction accuracy by minimizing the forecast error, we have taken the help of DWT. In this article, a discrete 1-level Haar wavelet transformation is performed to monthly electricity consumption time series data and the two components, wavelet approximation v1 and wavelet detailed component w1 are retrieved. Plots in fig. 3 depict the decomposed components.

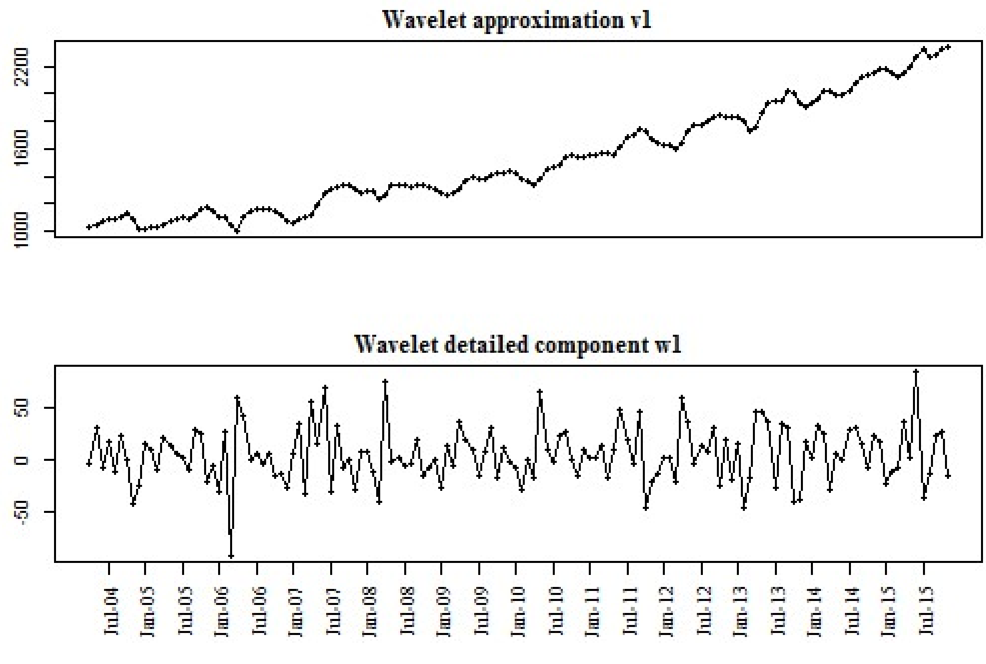

Figure 3: The 1-level Haar wavelet decomposition of monthly electricity consumption

To get the out-of-sample one-step forecasts we need to fit appropriate models to the components approximation v1 and detailed component w1. Training sets are required to fit appropriate models to both wavelet components and after the wavelet transform both components are split into two parts for both forecast horizons as described for electricity consumption. Sample ACF and sample PACF were used to assess the time series behaviour of both components. Also in case of training sets of wavelet approximation v1, persistent patterns of slowly declining ACF indicates lack of stationarity. To test the hidden seasonality, the chi-squared test has been applied to training sets of wavelet approximation v1 and the test results support the significance of seasonality with p-value 5.290213e-13 and 1.178861e-08 for both training sets and AICc values also show the significance of additional seasonality. An ETS model with additive trend and additive seasonal component ETS(A,A,A) model along with additive errors has been applied to the training set of wavelet approximation v1 with 110 observations and a model with additive seasonal component ETS(A,N,A) model including additive errors has been applied to the training set of wavelet approximation v1 with 93 observations. Also, based on the AICc values the model $ARIMA(1,1,2)(1,0,0)_{12}$ with drift was the best fit model for the training set of wavelet approximation v1 with 110 observations, indicating a first order non-seasonal difference and non-seasonal ARMA(1,2) and seasonal AR(1) components and a model $ARIMA(0,1,2)(2,0,0)_{12}$ was the best fit model for the training set of wavelet approximation v1 with 93 observations, indicating a first order non-seasonal difference and non-seasonal MA(2) and seasonal AR(2) components.

The ACF of training sets of detailed component w1 indicates the existence of stationarity. The chi-squared test results support the significance of seasonality with p-value 1.99885e-06 and 0.0001371544 for both training sets and AICc values also show the significance of additional seasonality. An ETS model with additive seasonal component ETS(A,N,A) model including additive errors has been applied to both training sets of detailed component w1. Also, based on the AICc values the model $ARIMA(1,0,0)(2,0,0)_{12}$ with zero mean was the best fit model for both training sets of detailed component w1, indicating a non-seasonal AR(1) and seasonal AR(2) components.

We summarize the results of the fitted univariate time series models for training sets of wavelet components used for prediction of electricity demand met in table 1 and table 2, to substantiate our findings. The summary provides the estimates of the coefficients along with standard errors and AICc values for fitted models.

Out-of-sample one-step prediction for the forecast horizon containing 31 observations with effect from June-2013 to December-2015 of training set of wavelet components are provided in the top panel of fig. 4 and out-of-sample one-step prediction for the forecast horizon containing 48 observations with effect from January-2012 to December-2015 of training set of wavelet components are provided in the top panel of fig. 5. The prediction performance results for both models and for both forecast horizons are tabulated in table 3. According to the performance results from table 3 and from the top panels of fig. 4 and fig. 5 shown above, it can be verified that SARIMA model gives better performance as compared to ETS model in case of training set of wavelet approximation v1 and ETS model gives better performance as compared to SARIMA in case of training set of wavelet component w1 in terms of performance accuracy measurement statistics mentioned above. Hence, combining SARIMA model forecasts for wavelet approximation v1 and ETS model forecasts for wavelet detailed component w1, the electricity consumption forecasts are computed.

Table 1: Estimated model coefficients along with standard errors and AICc values for fitted models (training set of 110 observations)

Model Description (training set of 110 observations)			
Series	**Model**	**Estimated parameter (Standard error)**	**AICc**
Training set of electricity consumption	ETS(A,A,A)	$\alpha = 0.5641$ $\beta = 0.0064$ $\gamma =$ 1e-04 $\sigma 2 = 1434.6065$	1354.453
	$ARIMA(1,1,0)(2,0,0)_{12}$	ar1 = -0.3017 (0.0943)*** sar1 = 0.3395 (0.0941)*** sar2 = 0.2415 (0.1080)** $\sigma 2 = 2042$	1152.7
Training set of approximation v1	ETS(A,A,A)	$\alpha = 0.9999$ $\beta =$ 1e-04 $\gamma =$ 1e-04 $\sigma 2 = 503.5356$	1227.43
	$ARIMA(1,1,2)(1,0,0)_{12}$ with drift	drift = 7.3379 (1.2976)*** ar1 = 0.7107 (0.1265)*** ma1 = -0.0078 (0.0959) ma2 = -0.8990 (0.0802)*** sar1 = 0.3992 (0.0929)*** $\sigma 2 = 522.9$	999.27
Training set of detailed component w1	ETS(A,N,A)	$\alpha =$ 2e-04 $\gamma =$ 1e-04 $\sigma 2 = 438.5380$	1206.92
	$ARIMA(1,0,0)(2,0,0)_{12}$ with zero mean	ar1 = -0.3017 (0.0943)*** sar1 = 0.3395 (0.0941)*** sar2 = 0.2415 (0.1080)** $\sigma 2 = 510.5$	1001.59

Notes: '***' and '**' denote significance at 1% and 5% level
Numbers in parenthesis are standard errors

Table 2: Estimated model coefficients along with standard errors and AICc values for fitted models (training set of 93 observations)

Model Description (training set of 93 observations)			
Series	**Model**	**Estimated parameter (Standard error)**	**AICc**
Training set of electricity consumption	ETS(A,N,A)	α = 0.6559 γ = 1e-04 σ2 = 1542.6435	1137.653
	$ARIMA(1,1,0)(2,0,0)_{12}$	ar1 = -0.3195 (0.1010)*** sar1 = 0.3443 (0.1036)*** sar2 = 0.2499 (0.1160)** σ2 = 1939	970.47
Training set of approximation v1	ETS(A,N,A)	α = 0.9998 γ = 1e-04 σ2 = 555.0783	1030.817
	$ARIMA(0,1,2)(2,0,0)_{12}$	ma1 = 0.6525 (0.1090)*** ma2 = -0.2726 (0.1160)** sar1 = 0.3300 (0.1043)*** sar2 = 0.2725 (0.1154)** σ2 = 497.3	839.89
Training set of detailed component w1	ETS(A,N,A)	α = 1e-04 γ = 1e-04 σ2 = 449.5757	1011.423
	$ARIMA(1,0,0)(2,0,0)_{12}$ with zero mean	ar1 = -0.3195 (0.1010)*** sar1 = 0.3443 (0.1036)*** sar2 = 0.2499 (0.1160)** σ2 = 484.7	842.93

Notes: '***' and '**' denote significance at 1% and 5% level
Numbers in parenthesis are standard errors

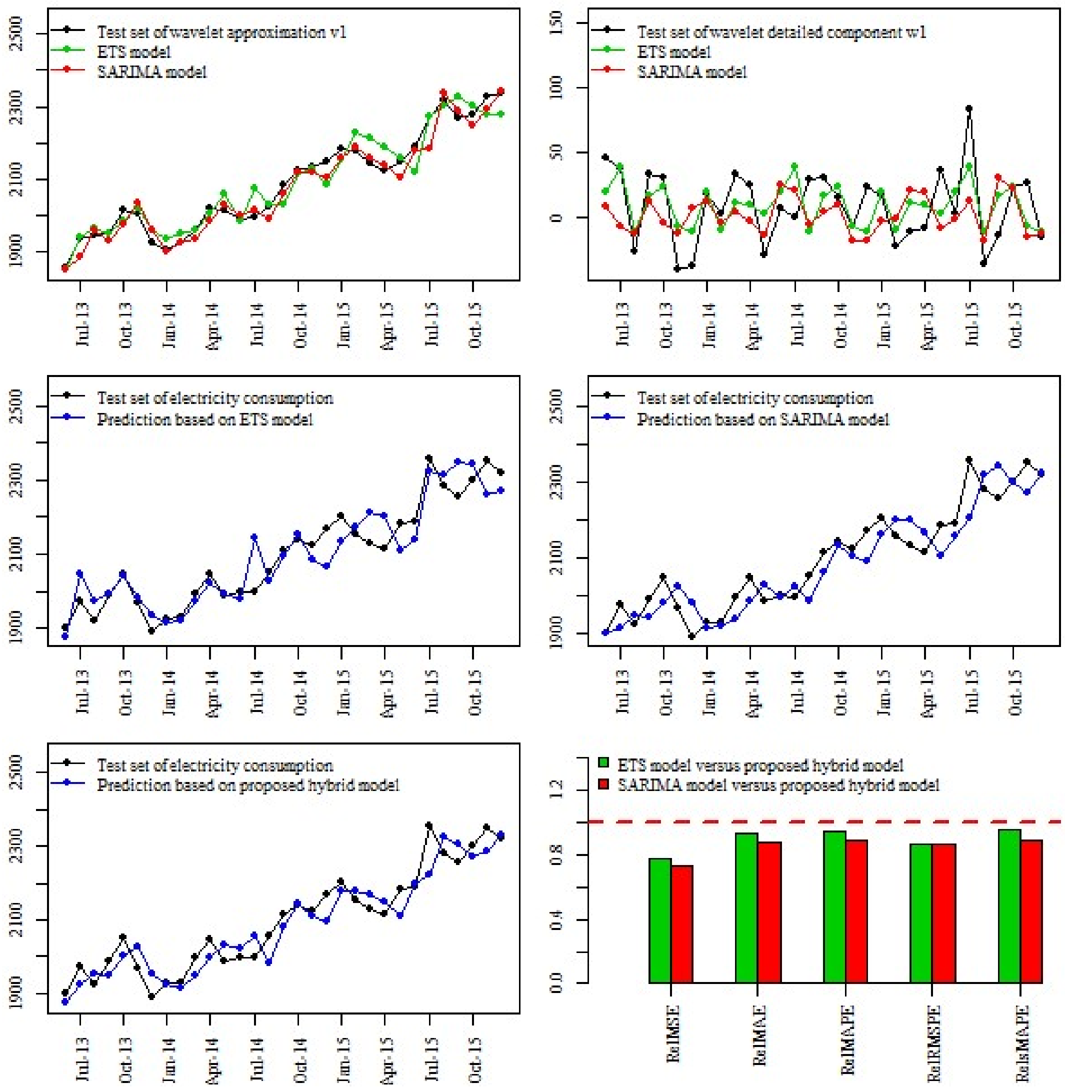

Figure 4: Out-of-sample one-step (31 points) prediction of training sets: v1 (approximation), w1 (detail) and performance comparison

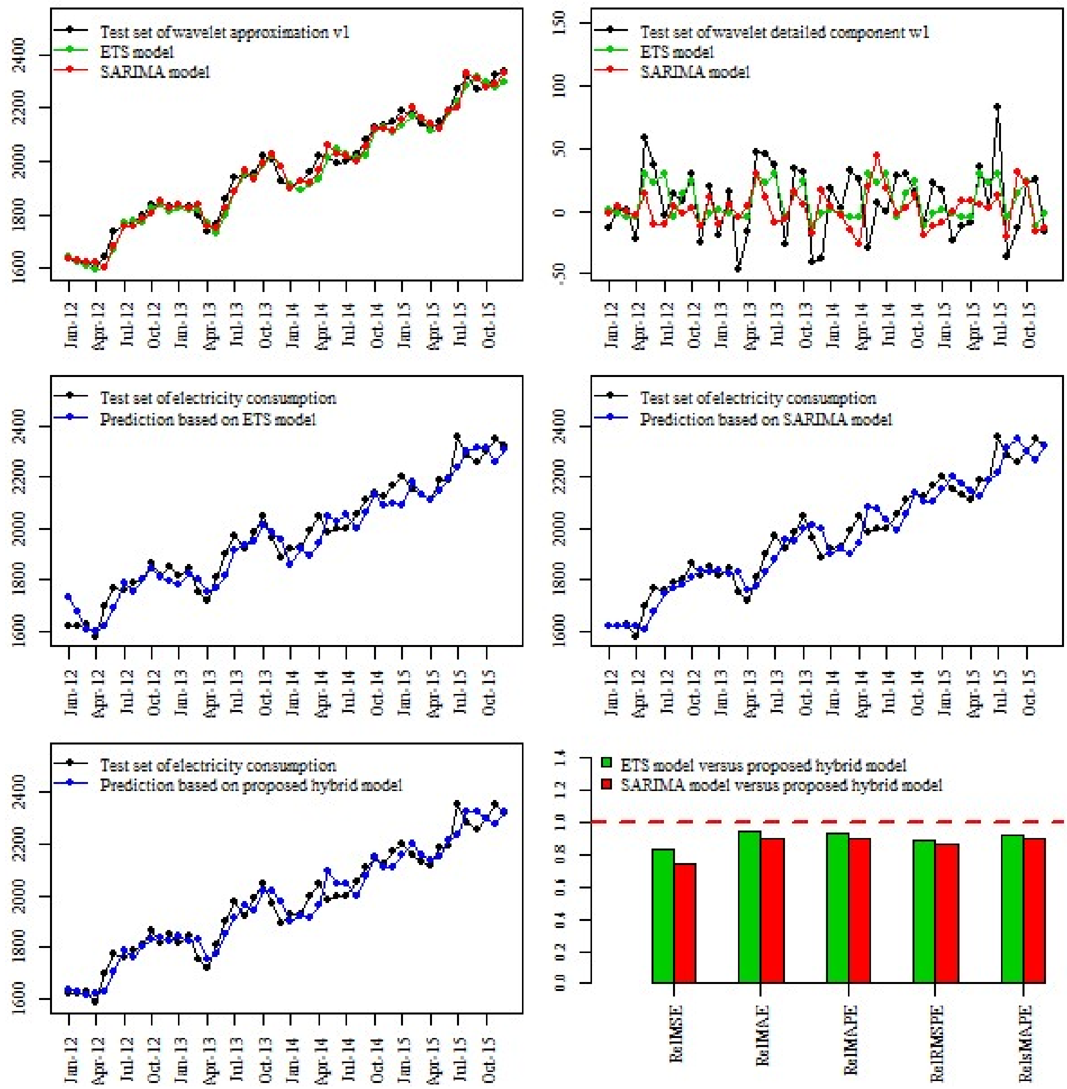

Figure 5: Out-of-sample one-step (48 points) prediction of training sets: v1 (approximation), w1 (detail) and performance comparison

Table 3: Commonly used forecast accuracy measures of performance comparison for wavelet approximation v1 and wavelet component w1

Measures	**Training set of wavelet approximation v1**		**Training set of wavelet detailed component w1**	
	ETS model	SARIMA model	ETS model	SARIMA model
31 points ahead				
MSE	1486.5943	865.7071	540.0115	905.2282
MAE	29.9884	23.8494	19.46439	25.56207
48 points ahead				
MSE	1161.8866	825.0087	560.9996	848.3006
MAE	27.01548	23.50267	19.8228	23.72435

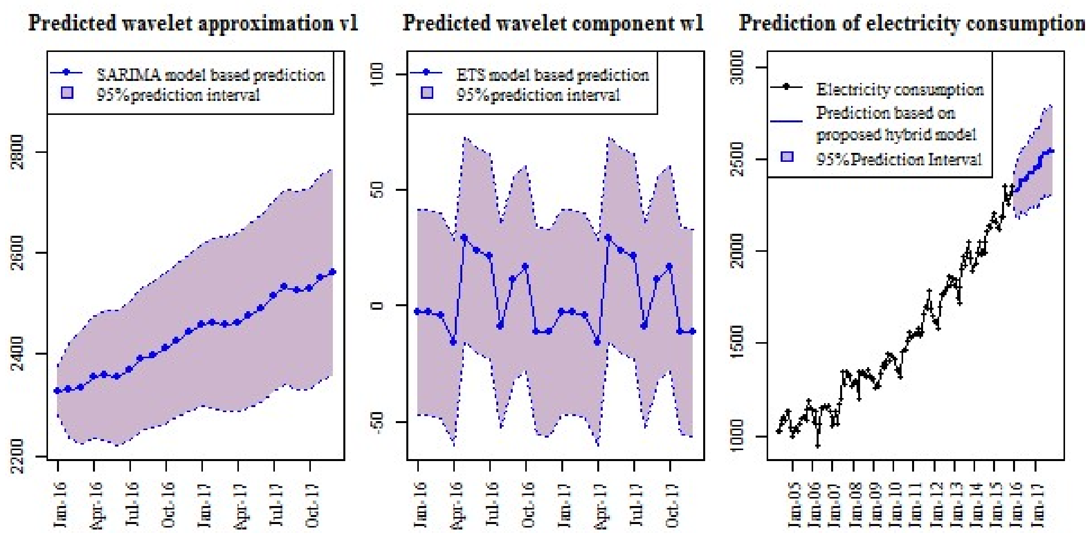

Figure 6: Predicted electricity consumption along with 95% prediction interval

Table 4: Commonly used forecast accuracy measures of performance comparison for electricity electricity consumption

Measures	**Training set of electricity consumption**			**Relative measures**	
	Prediction based on ETS model	Prediction based on SARIMA model	Prediction based on proposed hybrid model	ETS versus proposed hybrid model	SARIMA versus proposed hybrid model
31 points ahead					
MSE	3178.473	3330.789	2459.37	0.7737582	0.7383745
MAE	44.60641	47.53665	41.87297	0.938721	0.8808565
MAPE	2.103706	2.247989	1.989848	0.9458773	0.8851678
RMSPE	2.661961	2.690982	2.322423	0.8724482	0.863039
sMAPE	0.0209568	0.0226069	0.0199977	0.9542349	0.8845821
48 points ahead					
MSE	3047.566	3378.467	2526.745	0.8291028	0.747897
MAE	44.84478	46.92834	42.21026	0.9412525	0.8994621
MAPE	2.292139	2.362189	2.132662	0.9304244	0.902833
RMSPE	2.815711	2.905688	2.515313	0.8933139	0.8656517
sMAPE	0.023074	0.0237392	0.0213707	0.926181	0.9002289

Table 5: Estimated model coefficients along with standard error and AICc values for electricity consumption prediction

Model Description			
Series	**Model**	**Estimated parameter (Standard error)**	**AICc**
Wavelet approximation v1	$ARIMA(1,1,2)(0,0,2)_{12}$ with drift	drift = 9.0255 (1.9228)*** ar1 = 0.5640 (0.1127)*** ma1 = 0.1470 (0.0742)** ma2 = -0.8530 (0.0725)*** sma1 = 0.1619 (0.0902)* sma2 = 0.2491 (0.0889)*** $\sigma 2$ = 580.4	1299.92
Wavelet detailed component w1	ETS(A,N,A)	α = 1e-04 γ = 0.0015 $\sigma 2$ = 512.4926	1596.69

Table 6: Predicted electricity consumption based on proposed hybrid model along with 95% prediction interval

Forecast horizon	**Prediction based on proposed hybrid model along with 95% P.I.**	**Forecast horizon**	**Prediction based on proposed hybrid model along with 95% P.I.**
Jan-16	2326 (2234, 2417)	Jan-17	2455 (2251, 2659)
Feb-16	2327 (2189, 2465)	Feb-17	2458 (2247, 2669)
Mar-16	2331 (2176, 2486)	Mar-17	2455 (2240, 2671)
Apr-16	2339 (2174, 2504)	Apr-17	2447 (2227, 2667)
May-16	2388 (2216, 2559)	May-17	2505 (2280, 2729)
Jun-16	2377 (2200, 2554)	Jun-17	2513 (2285, 2741)
Jul-16	2390 (2209, 2571)	Jul-17	2535 (2303, 2767)
Aug-16	2380 (2195, 2565)	Aug-17	2523 (2288, 2759)
Sep-16	2411 (2222, 2600)	Sep-17	2537 (2297, 2776)
Oct-16	2428 (2236, 2621)	Oct-17	2547 (2304, 2790)
Nov-16	2417 (2221, 2613)	Nov-17	2539 (2293, 2785)
Dec-16	2432 (2232, 2631)	Dec-17	2551 (2302, 2801)

Bottom two panels of fig. 4 and fig. 5 depicts out-of-sample one-step prediction for the forecast horizon containing 31 and 48 observations with effect from June-2013 to December-2015 and with effect from January-2012 to December-2015 of training sets of electricity consumption using SARIMA model, ETS model and the proposed hybrid model. The prediction performance results for all three models and for both forecast horizons are tabulated in table 4. According to the performance results from table 4 and from fig. 4 and fig. 5, it can be verified that proposed hybrid model gives better performance as compared to SARIMA model and ETS model in terms of performance accuracy measurement statistics mentioned above.

With the forecast horizon of 31 points, the proposed hybrid model exhibits 23% and 13% more accurate forecasts in terms of MSE and RMSPE respectively than that of ETS model. On the other hand, the amount of increase in the accuracy of the proposed model are 26% and 14% in terms of MSE and RMSPE respectively in comparison to the SARIMA model. Further, with the forecast horizon of 48 points, the present model exhibits 17% and 11% more accurate forecasts in terms of MSE and RMSPE respectively than that of ETS model. While the amount of increase in the accuracy of the proposed model are 25% and 13% in terms of MSE and RMSPE respectively as compared to the SARIMA model. Hence it is evident that the robustness of the proposed model is adequately demonstrated in both the cases of medium and moderately long term forecasts.

Beyond the training set, two year forecasts with effect from January-2016 to December-2017 containing 24 points of wavelet approximation v1 using SARIMA model and wavelet detailed component w1 using ETS model are computed and provided in first two plot of fig. 6 along with the 95% prediction intervals. Model $ARIMA(1,1,2)(0,0,2)_{12}$ with drift was the best fit model for wavelet approximation v1, indicating a first order non-seasonal difference and non-seasonal ARMA(1,2) and seasonal MA(2) components. Wavelet detailed component w1 was fitted using an ETS model with additive seasonal component ETS(A,N,A) model along with additive errors. Model coefficients of fitted modes in both components are described in table 5 along with standard errors and AICc values. Combining the component wise forecasts, the electricity consumption forecasts with effect from January-2016 to December-2017 are computed and provided in the last plot of fig. 6 and also tabulated in table 6.

4 Concluding remarks

In this article, we have proposed a wavelet based hybrid model to forecast electricity consumption data from NER India. The purpose of this paper is to evaluate the superiority of wavelet based hybrid forecasting methods. Wavelets are used mainly in the context of data pre-processing. The actual forecast is done using one of the existing forecasting techniques, of which we have used two important family of modeling frameworks, viz., SARIMA and ETS. We also gave a brief introduction to the DWT used herein and then described how it is used in combination with SARIMA and ETS modeling.

The results of forecast accuracy for two different forecast horizons used in this study demonstrate that wavelet based forecasting method outperforms the individual conventional models like SARIMA or ETS as long as the data contain both linear and non-linear structures. It is also found that the model is reasonably robust in forecasting both for medium and moderately long term forecasts. A limitation of the present study is that we have not considered a method that suggests the choice of an optimal value of level n in DWT. However, that does not constrain us from using a wavelet based hybrid time series forecast model of the present kind.

Acknowledgement

The first author acknowledge the University Grant Commission for her financial support through Non-NET fellowship with award No.F.24-25/DSW/2012/1970

References

Afanasyev, D. O. and Fedorova, E. A. (2016). The long-term trends on the electricity markets: Comparison of empirical mode and wavelet decompositions. *Energy Economics*, 56:432–442.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6):716–723.

Bates, J. M. and Granger, C. W. (1969). The combination of forecasts. *Journal of the Operational Research Society*, 20(4):451–468.

Box, G. E. P. and Jenkins, G. M. (1970). *Time Series Analysis: Forecasting and Control.* Holden-Day San Francisco, California.

Box, G. E. P. and Jenkins, G. M. (1990). *Time Series Analysis: Forecasting and Control.* Holden-Day Incorporated.

Box, G. E. P., Jenkins, G. M., and Reinsel, G. C. (1994). *Time Series Analysis: Forecasting and Control.* Prentice Hall Press, Englewood Cliffs, NJ, third edition.

Brockwell, P. J. and Davis, R. A. (2006). *Introduction to time series and forecasting.* Springer Science and Business Media.

Brown, R. G. (1959). Statistical forecasting for inventory control.

Burnham, K. P. and Anderson, D. R. (2002). *Model Selection and Multi-model Inference: A Practical Information-Theoretic Approach.* Springer-Verlag, second edition.

Catalão, J. P. d. S., Pousinho, H. M. I., and Mendes, V. M. F. (2011). Hybrid wavelet-pso-anfis approach for short-term electricity prices forecasting. *IEEE Transactions on Power Systems*, 26(1):137–144.

Cavanaugh, J. E. (1997). Unifying the derivations for the akaike and corrected akaike information criteria. *Statistics & Probability Letters*, 33(2):201–208.

dos Santos Coelho, L. and Santos, A. A. (2011). A rbf neural network model with garch errors: application to electricity price forecasting. *Electric Power Systems Research*, 81(1):74–83.

Ediger, V. Ş. and Akar, S. (2007). Arima forecasting of primary energy demand by fuel in turkey. *Energy Policy*, 35(3):1701–1708.

Egelioglu, F., Mohamad, A., and Guven, H. (2001). Economic variables and electricity consumption in northern cyprus. *Energy*, 26(4):355–362.

Erdogdu, E. (2007). Electricity demand analysis using cointegration and arima modelling: A case study of turkey. *Energy policy*, 35(2):1129–1146.

Garcia, R. C., Contreras, J., Van Akkeren, M., and Garcia, J. B. C. (2005). A garch forecasting model to predict day-ahead electricity prices. *IEEE transactions on power systems*, 20(2):867–874.

Gençay, R., Selçuk, F., and Whitcher, B. J. (2001). *An introduction to wavelets and other filtering methods in finance and economics.* Academic press.

Granger, C. W. J. and Newbold, P. (2014). *Forecasting economic time series.* Academic Press.

Haar, A. (1910). Zur theorie der orthogonalen funktionensysteme. *Mathematische Annalen*, 69(3):331–371.

Holt, C. C. (1957). Forecasting trends and seasonals by exponentially weighted averages.

Hyndman, R., Koehler, A. B., Ord, J. K., and Snyder, R. D. (2008). *Forecasting with exponential smoothing: the state space approach.* Springer Science & Business Media.

Hyndman, R. J., Athanasopoulos, G., Razbash, S., Schmidt, D., Zhou, Z., Khan, Y., Bergmeir, C., and Wang, E. (2014). forecast: Forecasting functions for time series and linear models. *R package version 5.5.*

Hyndman, R. J. and Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International journal of forecasting*, 22(4):679–688.

Hyndman, R. J., Koehler, A. B., Snyder, R. D., and Grose, S. (2002). A state space framework for automatic forecasting using exponential smoothing methods. *International Journal of Forecasting*, 18(3):439–454.

Liu, H. and Shi, J. (2013). Applying arma–garch approaches to forecasting short-term electricity prices. *Energy Economics*, 37:152–166.

Meitz, M. and Saikkonen, P. (2013). Maximum likelihood estimation of a noninvertible arma model with autoregressive conditional heteroskedasticity. *Journal of Multivariate Analysis*, 114:227–255.

Mohamed, Z. and Bodger, P. (2003). Analysis of the logistic model for predicting new zealand electricity consumption.

Mohamed, Z. and Bodger, P. (2005). A comparison of logistic and harvey models for electricity consumption in new zealand. *Technological Forecasting and Social Change*, 72(8):1030–1043.

North-Eastern Regional Power Committee (NERPC) (2015). *Annual Reports, 2005-06; 2007-08; 2009-10; 2011-12; 2012-13 and 2014-15.* Ministry of Power, Government of India, Shillong-793006, Meghalaya.

R Development Core Team (2008). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-

07-0.

Schwarz, G. et al. (1978). Estimating the dimension of a model. *The annals of statistics*, 6(2):461–464.

Shrivastava, N. A. and Panigrahi, B. K. (2014). A hybrid wavelet-elm based short term price forecasting for electricity markets. *International Journal of Electrical Power & Energy Systems*, 55:41–50.

Sims, C. A. et al. (1986). Are forecasting models usable for policy analysis? *Federal Reserve Bank of Minneapolis Quarterly Review*, 10(1):2–16.

Tan, Z., Zhang, J., Wang, J., and Xu, J. (2010). Day-ahead electricity price forecasting using wavelet transform combined with arima and garch models. *Applied Energy*, 87(11):3606–3610.

Taylor, J. W. (2003). Short-term electricity demand forecasting using double seasonal exponential smoothing. *Journal of the Operational Research Society*, 54(8):799–805.

Walker James, S. (1999). A primer on wavelets and their scientific applications [livro].-[sl].

Wang, J., Liu, F., Song, Y., and Zhao, J. (2016). A novel model: Dynamic choice artificial neural network (dcann) for an electricity price forecasting system. *Applied Soft Computing*, 48:281–297.

Wang, Y., Wang, J., Zhao, G., and Dong, Y. (2012). Application of residual modification approach in seasonal arima for electricity demand forecasting: a case study of china. *Energy Policy*, 48:284–294.

Wheelwright, S., Makridakis, S., and Hyndman, R. J. (1998). *Forecasting: methods and applications*. John Wiley & Sons.

Whitcher, B. (2015). Waveslim: Basic wavelet routines for one-, two- and three-dimensional signal processing. *R package version 1.7.5*.

Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages. *Management Science*, 6(3):324–342.

Yan, X. and Chowdhury, N. A. (2013). Mid-term electricity market clearing price forecasting: A hybrid lssvm and armax approach. *International Journal of Electrical Power & Energy Systems*, 53:20–26.

Yan, Y. Y. (1998). Climate and residential electricity consumption in hong kong. *Energy*, 23(1):17–20.

Zahan, M. and Kenett, R. S. (2013). Modeling and forecasting energy consumption in the manufacturing industry in south asia. *International Journal of Energy Economics and Policy*, 3(1):87.

Zhang, J. and Tan, Z. (2013). Day-ahead electricity price forecasting using wt, clssvm and egarch model. *International Journal of Electrical Power & Energy Systems*, 45(1):362–368.